



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Moore, David
Seol, Wongi
Choi, Hueng-Sik
- (ii) TITLE OF INVENTION: RETINOID X RECEPTOR-INTERACTING
POLYPEPTIDES AND RELATED MOLECULES AND METHODS
- (iii) NUMBER OF SEQUENCES: 17
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Fish & Richardson P.C.
 - (B) STREET: 225 Franklin Street, Suite 3100
 - (C) CITY: Boston
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02110-2804
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/372,652
 - (B) FILING DATE: 13-JAN-1995
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Clark, Paul T.
 - (B) REGISTRATION NUMBER: 30,162
 - (C) REFERENCE/DOCKET NUMBER: 00786/246001
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 617/542-5070
 - (B) TELEFAX: 617/542-8906
 - (C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 484 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: not relevant
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met Val Met Gln Phe Gln Gly Leu Glu Asn Pro Ile Gln Ile Ser Leu
1 5 10 15
His His Ser His Arg Leu Ser Gly Phe Val Pro Asp Gly Met Ser Val
20 25 30
Lys Pro Ala Lys Gly Met Leu Thr Glu His Ala Ala Gly Pro Leu Gly
35 40 45
Gln Asn Leu Asp Leu Glu Ser Tyr Ser Pro Tyr Asn Asn Val Pro Phe
50 55 60
Pro Gln Val Gln Pro Gln Ile Ser Ser Ser Ser Tyr Tyr Ser Asn Leu
65 70 75 80
Gly Phe Tyr Pro Gln Gln Pro Glu Asp Trp Tyr Ser Pro Gly Ile Tyr
85 90 95
Glu Leu Arg Arg Met Pro Ala Glu Thr Gly Tyr Gln Gly Glu Thr Glu
100 105 110
Val Ser Glu Met Pro Val Thr Lys Lys Pro Arg Met Ala Ala Ala Ser
115 120 125
Ala Gly Arg Ile Lys Gly Asp Glu Leu Cys Val Val Cys Gly Asp Arg
130 135 140
Ala Ser Gly Tyr His Tyr Asn Ala Leu Thr Cys Glu Gly Cys Lys Gly
145 150 155 160
Phe Phe Arg Arg Ser Ile Thr Lys Asn Ala Val Tyr Lys Cys Lys Asn
165 170 175
Gly Gly Asn Cys Val Met Asp Met Tyr Met Arg Arg Lys Cys Gln Glu
180 185 190
Cys Arg Leu Arg Lys Cys Arg Glu Met Gly Met Leu Ala Glu Cys Leu
195 200 205
Leu Thr Glu Ile Gln Cys Lys Ser Lys Arg Leu Arg Lys Asn Val Lys
210 215 220
Gln His Ala Asp Gln Thr Val Asn Glu Asp Asp Ser Glu Gly Arg Asp
225 230 235 240
Leu Arg Gln Val Thr Ser Thr Thr Lys Phe Cys Arg Glu Lys Thr Glu
245 250 255
Leu Thr Ala Asp Gln Gln Thr Leu Leu Asp Tyr Ile Met Asp Ser Tyr
260 265 270
Asn Lys Gln Arg Met Pro Gln Glu Ile Thr Asn Lys Ile Leu Lys Glu
275 280 285
Glu Phe Ser Ala Glu Glu Asn Phe Leu Ile Leu Thr Glu Met Ala Thr
290 295 300
Ser His Val Gln Ile Leu Val Glu Phe Thr Lys Lys Leu Pro Gly Phe
305 310 315 320

JAN 11 1960

Gln Thr Leu Asp His Glu Asp Gln Ile Ala Leu Leu Lys Gly Ser Ala
 325 330 335
 Val Glu Ala Met Phe Leu Arg Ser Ala Glu Ile Phe Asn Lys Lys Leu
 340 345 350
 Pro Ala Gly His Ala Asp Leu Leu Glu Glu Arg Ile Arg Lys Ser Gly
 355 360 365
 Ile Ser Asp Glu Tyr Ile Thr Pro Met Phe Ser Phe Tyr Lys Ser Val
 370 375 380
 Gly Glu Leu Lys Met Thr Gln Glu Glu Tyr Ala Leu Leu Thr Ala Ile
 385 390 395 400
 Val Ile Leu Ser Pro Asp Arg Gln Tyr Ile Lys Asp Arg Glu Ala Val
 405 410 415
 Glu Lys Leu Gln Glu Pro Leu Leu Asp Val Leu Gln Lys Leu Cys Lys
 420 425 430
 Met Tyr Gln Pro Glu Asn Pro Gln His Phe Ala Cys Leu Leu Gly Arg
 435 440 445
 Leu Thr Glu Leu Arg Thr Phe Asn His His His Ala Glu Met Leu Met
 450 455 460
 Ser Trp Arg Val Asn Asp His Lys Phe Thr Pro Leu Leu Cys Glu Ile
 465 470 475 480
 Trp Asp Val Gln

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 451 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Thr Glu His Ala Ala Gly Pro Leu Gly Gln Asn Leu Asp Leu
 1 5 10 15
 Glu Ser Tyr Ser Pro Tyr Asn Asn Val Pro Phe Pro Gln Val Gln Pro
 20 25 30
 Gln Ile Ser Ser Ser Tyr Tyr Ser Asn Leu Gly Phe Tyr Pro Gln
 35 40 45
 Gln Pro Glu Asp Trp Tyr Ser Pro Gly Ile Tyr Glu Leu Arg Arg Met
 50 55 60
 Pro Ala Glu Thr Gly Tyr Gln Gly Glu Thr Glu Val Ser Glu Met Pro
 65 70 75 80

Val	Thr	Lys	Lys	Pro	Arg	Met	Ala	Ala	Ala	Ser	Ala	Gly	Arg	Ile	Lys		
				85					90					95			
Gly	Asp	Glu	Leu	Cys	Val	Val	Cys	Gly	Asp	Arg	Ala	Ser	Gly	Tyr	His		
			100					105					110				
Tyr	Asn	Ala	Leu	Thr	Cys	Glu	Gly	Cys	Lys	Gly	Phe	Phe	Arg	Arg	Ser		
		115					120					125					
Ile	Thr	Lys	Asn	Ala	Val	Tyr	Lys	Cys	Lys	Asn	Gly	Gly	Asn	Cys	Val		
		130				135					140						
Met	Asp	Met	Tyr	Met	Arg	Arg	Lys	Cys	Gln	Glu	Cys	Arg	Leu	Arg	Lys		
		145			150					155					160		
Cys	Arg	Glu	Met	Gly	Met	Leu	Ala	Glu	Cys	Met	Tyr	Thr	Gly	Leu	Leu		
			165					170						175			
Thr	Glu	Ile	Gln	Cys	Lys	Ser	Lys	Arg	Leu	Arg	Lys	Asn	Val	Lys	Gln		
			180					185					190				
His	Ala	Asp	Gln	Thr	Val	Asn	Glu	Asp	Asp	Ser	Glu	Gly	Arg	Asp	Leu		
		195					200					205					
Arg	Gln	Val	Thr	Ser	Thr	Thr	Lys	Phe	Cys	Arg	Glu	Lys	Thr	Glu	Leu		
		210				215					220						
Thr	Ala	Asp	Gln	Gln	Thr	Leu	Leu	Asp	Tyr	Ile	Met	Asp	Ser	Tyr	Asn		
		225			230				235						240		
Lys	Gln	Arg	Met	Pro	Gln	Glu	Ile	Thr	Asn	Lys	Ile	Leu	Lys	Glu	Glu		
			245						250					255			
Phe	Ser	Ala	Glu	Glu	Asn	Phe	Leu	Ile	Leu	Thr	Glu	Met	Ala	Thr	Ser		
		260					265						270				
His	Val	Gln	Ile	Leu	Val	Glu	Phe	Thr	Lys	Lys	Leu	Pro	Gly	Phe	Gln		
		275					280					285					
Thr	Leu	Asp	His	Glu	Asp	Gln	Ile	Ala	Leu	Leu	Lys	Gly	Ser	Ala	Val		
		290				295					300						
Glu	Ala	Met	Phe	Leu	Arg	Ser	Ala	Glu	Ile	Phe	Asn	Lys	Lys	Leu	Pro		
		305			310					315					320		
Ala	Gly	His	Ala	Asp	Leu	Leu	Glu	Glu	Arg	Ile	Arg	Lys	Ser	Gly	Ile		
			325						330					335			
Ser	Asp	Glu	Tyr	Ile	Thr	Pro	Met	Phe	Ser	Phe	Tyr	Lys	Ser	Val	Gly		
		340						345					350				
Glu	Leu	Lys	Met	Thr	Gln	Glu	Glu	Tyr	Ala	Leu	Leu	Thr	Ala	Ile	Val		
		355					360					365					
Ile	Leu	Ser	Pro	Asp	Arg	Gln	Tyr	Ile	Lys	Asp	Arg	Glu	Ala	Val	Glu		
		370				375					380						
Lys	Leu	Gln	Glu	Pro	Leu	Leu	Asp	Val	Leu	Gln	Lys	Leu	Cys	Lys	Met		
		385			390					395					400		
Tyr	Gln	Pro	Glu	Asn	Pro	Gln	His	Phe	Ala	Cys	Leu	Leu	Gly	Arg	Leu		
			405						410					415			

Thr Glu Leu Arg Thr Phe Asn His His His Ala Glu Met Leu Met Ser
420 425 430

Trp Arg Val Asn Asp His Lys Phe Thr Pro Leu Leu Cys Glu Ile Trp
435 440 445

Asp Val Gln
450

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 446 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Ser Ser Pro Thr Ser Ser Leu Asp Thr Pro Val Pro Gly Asn Gly
1 5 10 15

Ser Pro Gln Pro Ser Thr Ser Ala Thr Ser Pro Thr Ile Lys Glu Glu
20 25 30

Gly Gln Glu Thr Asp Pro Pro Pro Gly Ser Glu Gly Ser Ser Ser Ala
35 40 45

Tyr Ile Val Val Ile Leu Glu Pro Glu Asp Glu Pro Glu Arg Lys Arg
50 55 60

Lys Lys Gly Pro Ala Pro Lys Met Leu Gly His Glu Leu Cys Arg Val
65 70 75 80

Cys Gly Asp Lys Ala Ser Gly Phe His Tyr Asn Val Leu Ser Cys Glu
85 90 95

Gly Cys Lys Gly Phe Phe Arg Arg Ser Val Val His Gly Gly Ala Gly
100 105 110

Arg Tyr Ala Cys Arg Gly Ser Gly Thr Cys Gln Met Asp Ala Phe Met
115 120 125

Arg Arg Lys Cys Gln Leu Cys Arg Leu Arg Lys Cys Lys Glu Ala Gly
130 135 140

Met Arg Glu Gln Cys Val Leu Ser Glu Glu Gln Ile Arg Lys Lys Arg
145 150 155 160

Ile Gln Lys Gln Gln Gln Gln Gln Pro Pro Pro Pro Ser Glu Pro Ala
165 170 175

Ala Ser Ser Ser Gly Arg Pro Ala Ala Ser Pro Gly Thr Ser Glu Ala
180 185 190

Ser Ser Gln Gly Ser Gly Glu Gly Glu Gly Ile Gln Leu Thr Ala Ala
195 200 205

Gln	Glu	Leu	Met	Ile	Gln	Gln	Leu	Val	Ala	Ala	Gln	Leu	Gln	Cys	Asn
210					215						220				
Lys	Arg	Ser	Phe	Ser	Asp	Gln	Pro	Lys	Val	Thr	Pro	Trp	Pro	Leu	Gly
225					230					235					240
Ala	Asp	Pro	Gln	Ser	Arg	Asp	Ala	Arg	Gln	Gln	Arg	Phe	Ala	His	Phe
				245					250					255	
Thr	Glu	Leu	Ala	Ile	Ile	Ser	Val	Gln	Glu	Ile	Val	Asp	Phe	Ala	Lys
			260					265					270		
Gln	Val	Pro	Gly	Phe	Leu	Gln	Leu	Gly	Arg	Glu	Asp	Gln	Ile	Ala	Leu
		275					280					285			
Leu	Lys	Ala	Ser	Thr	Ile	Glu	Ile	Met	Leu	Leu	Gln	Thr	Ala	Arg	Arg
	290					295					300				
Tyr	Asn	His	Glu	Thr	Glu	Cys	Ile	Thr	Phe	Leu	Lys	Asp	Phe	Thr	Tyr
305					310					315					320
Ser	Lys	Asp	Asp	Phe	His	Arg	Ala	Gly	Leu	Gln	Val	Glu	Phe	Ile	Asn
				325					330					335	
Pro	Ile	Phe	Glu	Phe	Ser	Arg	Ala	Met	Arg	Arg	Leu	Gly	Leu	Asp	Asp
			340					345					350		
Ala	Glu	Tyr	Ala	Leu	Leu	Ile	Ala	Ile	Asn	Ile	Phe	Ser	Ala	Asp	Arg
		355					360					365			
Pro	Asn	Val	Gln	Glu	Pro	Ser	Arg	Val	Glu	Ala	Leu	Gln	Gln	Pro	Tyr
	370					375					380				
Val	Glu	Ala	Leu	Leu	Ser	Tyr	Thr	Arg	Ile	Lys	Arg	Pro	Gln	Asp	Gln
385					390					395					400
Leu	Arg	Phe	Pro	Arg	Met	Leu	Met	Lys	Leu	Val	Ser	Leu	Arg	Thr	Leu
				405					410					415	
Ser	Ser	Val	His	Ser	Glu	Gln	Val	Phe	Ala	Leu	Arg	Leu	Gln	Asp	Lys
			420					425					430		
Lys	Leu	Pro	Pro	Leu	Leu	Ser	Glu	Ile	Trp	Asp	Val	His	Glu		
		435					440					445			

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 716 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ser Pro Ser Cys Gly Phe Gln Arg Lys Asp Leu Glu Thr Ser Ser Cys
1 5 10 15

Val Ser Ile Lys Lys Lys Arg Arg Leu Glu Asp Leu Leu Ile Val Ile
20 25 30

Ser Asp Ser Asp Gly Glu Glu Thr Lys Glu Glu Asn Gly Leu Gln Lys
35 40 45

Thr Lys Thr Lys Gln Ser Asn Arg Ser Lys Cys Leu Ala Lys Arg Lys
50 55 60

Val Ala His Met Ser Glu Glu Gln Phe Ala Leu Ala Leu Lys Met
65 70 75 80

Ser Glu Gln Glu Ala Arg Glu Val Asn Asn Gln Glu Glu Lys Glu Glu
85 90 95

Glu Leu Leu Arg Lys Ala Ile Ala Glu Ser Leu Asn Ser Cys Trp Ser
100 105 110

Ser Ala Ala Ser Ala Thr Arg Ser Arg Pro Leu Ala Ala Glu Leu Ser
115 120 125

Ser His Ser His Gln Glu Asn Thr Lys Asp Ser Gly Thr Thr Glu Gly
130 135 140

Val Trp Gln Leu Val Pro Pro Ser Leu Cys Lys Gly Ser His Val Ser
145 150 155 160

Gln Gly Asn Glu Ala Glu Gln Arg Lys Glu Pro Trp Asp His Asn Glu
165 170 175

Asn Thr Glu Glu Glu Pro Val Ser Gly Ser Ser Gly Ser Trp Asp Gln
180 185 190

Ser Ser Gln Pro Val Phe Glu Asn Glu Asn Val Lys Cys Phe Asp Arg
195 200 205

Cys Thr Gly His Leu Ala Glu His Thr Gln Cys Gly Lys Pro Gln Glu
210 215 220

Ser Thr Gly Ser Gly Tyr Ala Phe Ser Lys Ala Val Gln Gly Arg Gly
225 230 235 240

Asp Thr Ser Arg Gln Cys Leu Pro Ile Pro Ala Asp Thr Lys Gly Leu
245 250 255

Gln Asp Thr Gly Gly Thr Val His Tyr Tyr Trp Gly Ile Pro Phe Cys
260 265 270

Pro Ala Gly Val Asp Pro Asn Gln Tyr Thr Asn Val Ile Leu Cys Gln
275 280 285

Leu Glu Val Tyr Gln Lys Ser Leu Lys Met Ala Gln Arg Gln Leu Val
290 295 300

Lys Lys Arg Gly Phe Gly Glu Pro Val Leu Pro Arg Pro Pro Phe Leu
305 310 315 320

Ile Gln Asn Glu Cys Gly Gln Glu Asp Gln Thr Ser Asp Lys Asn Glu
 325 330 335
 Gly Ile Ser Glu Asp Met Gly Asp Glu Ala Lys Glu Glu Arg Gln Glu
 340 345 350
 Ser Arg Ala Ser Val Trp His Ser Glu Thr Lys Asp Phe Gln Lys Ser
 355 360 365
 Pro Ile Lys Ser Leu Lys Gln Lys Leu Leu Leu Glu Glu Glu Pro Thr
 370 375 380
 Thr Ser Arg Gly Gln Ser Ser Gln Gly Leu Phe Val Glu Glu Thr Ser
 385 390 395 400
 Glu Glu Gly Leu Lys Ser Ser Glu Gly Asp Asn Ser Val Pro Thr Thr
 405 410 415
 Gln Ser Ile Ala Ala Leu Thr Ser Lys Arg Ser Leu Val Leu Met Pro
 420 425 430
 Glu Ser Ser Ala Glu Glu Ile Thr Val Cys Pro Glu Thr Gln Leu Ser
 435 440 445
 Phe Leu Glu Pro Leu Asp Leu Asn Arg Glu Asp Ser Pro Asp Ser Arg
 450 455 460
 Glu Leu Pro Ile Glu Val Arg Met Ala Val Gly Asp Lys Gln Val Ala
 465 470 475 480
 Asn Arg Glu Asp Cys Met Lys Glu Asn Pro Pro Pro Ala Val Ser Ser
 485 490 495
 Ser Thr Arg Val Ser Cys Pro Leu Cys Asn Gln Asp Phe Pro Pro Thr
 500 505 510
 Lys Ile Glu Gln His Ala Met Tyr Cys Asn Gly Leu Met Glu Gln Glu
 515 520 525
 Thr Val Leu Thr Arg Arg Arg Arg Glu Ala Lys Asn Lys Ser Asp Gly
 530 535 540
 Arg Thr Ala Ala Gln Pro Ala Leu Asp Ala Asn Arg Lys Glu Lys Cys
 545 550 555 560
 Tyr Leu Cys Lys Ser Leu Val Pro Leu Gly Glu Tyr Gln Cys His Val
 565 570 575
 Glu Ala Cys Leu Gln Leu Ala Lys Val Asp Arg Glu Asp Gly Ile Glu
 580 585 590
 Gly Thr Arg Arg Pro Arg Val Cys Ala Pro Val Glu Gly Lys Gln Gln
 595 600 605
 Gln Arg Leu Lys Lys Ser Lys Asp Lys Gly His Ser Gln Gly Arg Leu
 610 615 620
 Leu Ser Leu Leu Glu Gln Ser Glu His Arg Thr Thr Gly Val Glu Lys
 625 630 635 640
 Lys Pro Lys Tyr Ser Glu Val Arg Thr Phe Arg Met Pro Ser Pro Glu
 645 650 655

Val Glu Glu Ala Ser Cys Ser Arg Glu Met Gln Ser Thr Leu Ser Gln
660 665 670

Leu Asn Leu Asn Glu Ser Pro Ile Lys Ser Phe Val Pro Val Ser Glu
675 680 685

Ala Thr Asn Cys Leu Val Asp Phe Lys Glu Gln Phe Ala Phe Arg Ser
690 695 700

Arg Thr Lys Ser Gly Arg Glu Arg Arg Arg Lys Ser
705 710 715

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 619 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Ala Leu Ala Ala Leu Val Asp Ala Ala Ala Ser Ala Pro Gln Met
1 5 10 15

Asp Val Ser Lys Thr Lys Glu Ser Lys His Glu Ala Ala Arg Leu Glu
20 25 30

Glu Asn Leu Arg Ser Arg Ser Ala Ala Val Ser Glu Gln Gln Gln Leu
35 40 45

Glu Gln Lys Asn Leu Glu Val Glu Lys Arg Ser Val Gln Cys Val Cys
50 55 60

Thr Ser Ser Ala Leu Pro Ser Gly Lys Ala Gln Pro His Ala Ser Val
65 70 75 80

Val Tyr Ser Glu Ala Gly Lys Asp Lys Gly Pro Pro Pro Lys Ser Arg
85 90 95

Tyr Glu Glu Glu Leu Arg Thr Arg Gly Lys Thr Thr Ile Thr Ala Ala
100 105 110

Asn Phe Ile Asp Val Thr Ile Thr Arg Gln Ile Ala Ser Asp Lys Asp
115 120 125

Ala Arg Glu Arg Gly Ser Gln Ser Ser Asp Ser Ser Ser Ser Leu Ser
130 135 140

Ser His Arg Tyr Glu Thr Ala Ser Asp Ala Ile Glu Val Ile Ser Pro
145 150 155 160

Ala Ser Ser Pro Ala Pro Pro Gln Glu Lys Pro Gln Ala Tyr Gln Pro
165 170 175

Asp Met Val Lys Ala Asn Gln Ala Glu Asn Glu Ser Thr Arg Gln Tyr
180 185 190

Glu Gly Pro Leu His His Tyr Arg Ser Gln Gln Glu Ser Pro Ser Pro
 195 200 205
 Gln Gln Gln Pro Pro Leu Pro Pro Ser Ser Gln Ser Glu Gly Met Gly
 210 215 220
 Gln Val Pro Arg Thr His Arg Leu Ile Thr Leu Ala Asp His Ile Cys
 225 230 235 240
 Gln Ile Ile Thr Gln Asp Phe Ala Arg Asn Gln Val Pro Ser Gln Pro
 245 250 255
 Ser Thr Ser Thr Phe Gln Thr Ser Pro Ser Ala Leu Ser Ser Thr Pro
 260 265 270
 Val Arg Thr Lys Thr Ser Ser Arg Tyr Ser Pro Glu Ser Gln Ser Gln
 275 280 285
 Thr Val Leu His Pro Arg Pro Gly Pro Arg Val Ser Pro Glu Asn Leu
 290 295 300
 Val Asp Lys Ser Arg Gly Ser Arg Pro Gly Lys Ser Pro Glu Arg Ser
 305 310 315 320
 His Ile Pro Ser Glu Pro Tyr Glu Pro Ile Ser Pro Pro Gln Gly Pro
 325 330 335
 Ala Val His Glu Lys Gln Asp Ser Met Leu Leu Leu Ser Gln Arg Gly
 340 345 350
 Val Asp Pro Ala Glu Gln Arg Ser Asp Ser Arg Ser Pro Gly Ser Ile
 355 360 365
 Ser Tyr Leu Pro Ser Phe Phe Thr Lys Leu Glu Ser Thr Ser Pro Met
 370 375 380
 Val Lys Ser Lys Lys Gln Glu Ile Phe Arg Lys Leu Asn Ser Ser Gly
 385 390 395 400
 Gly Gly Asp Ser Asp Met Ala Ala Ala Gln Pro Gly Thr Glu Ile Phe
 405 410 415
 Asn Leu Pro Ala Val Thr Thr Ser Gly Ala Val Ser Ser Arg Ser His
 420 425 430
 Ser Phe Ala Asp Pro Ala Ser Asn Leu Gly Leu Glu Asp Ile Ile Arg
 435 440 445
 Lys Ala Leu Met Gly Ser Phe Asp Asp Lys Val Glu Asp His Gly Val
 450 455 460
 Val Met Ser His Pro Val Gly Ile Met Pro Gly Ser Ala Ser Thr Ser
 465 470 475 480
 Val Val Thr Ser Ser Glu Ala Arg Arg Asp Glu Gly Glu Pro Ser Pro
 485 490 495
 His Ala Gly Val Cys Lys Pro Lys Leu Ile Asn Lys Ser Asn Ser Arg
 500 505 510
 Lys Ser Lys Ser Pro Ile Pro Gly Gln Ser Tyr Leu Gly Thr Glu Arg
 515 520 525

Pro Ser Ser Val Ser Ser Val His Ser Glu Gly Asp Tyr His Arg Gln
 530 535 540
 Thr Pro Gly Trp Ala Trp Glu Asp Arg Pro Ser Ser Thr Gly Ser Thr
 545 550 555 560
 Gln Phe Pro Tyr Asn Pro Leu Thr Ile Arg Met Leu Ser Ser Thr Pro
 565 570 575
 Pro Thr Gln Ile Ala Cys Ala Pro Ser Ala Ile Thr Gln Ala Ala Pro
 580 585 590
 His Gln Gln Asn Arg Ile Trp Glu Arg Glu Pro Ala Pro Leu Leu Ser
 595 600 605
 Ala Gln Tyr Glu Thr Leu Ser Asp Ser Asp Asp
 610 615

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1787 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGAAGCTAAG GATGGTGATG CAGTTTCAGG GCTTAGAAAA TCCAATTCAG ATTAGTCTTC	60
ACCACAGCCA CCGGCTGTCA GGATTTGTGC CGGACGGGAT GAGTGTGAAG CCAGCTAAAG	120
GTATGCTAAC AGAACACGCG GCAGGCCCTC TGGGGCAGAA TCTGGATTG GAATCGTACT	180
CCCCATACAA CAATGTCCCG TTTCCTCAAG TTCAGCCACA GATTTCTCTC TCGTCTTACT	240
ATCCAACCT GGGCTTCTAC CCCCACAAC CGGAAGACTG GTATTCTCCT GGCATCTATG	300
AACTCAGGCG AATGCCCCGCT GAGACTGGGT ACCAGGGAGA GACTGAGGTA TCAGAGATGC	360
CTGTGACAAA GAAGCCGCGA ATGGCCGCGG CATCGGCAGG CAGAATAAAA GGGGATGAGC	420
TGTGTGTTGT CTGTGGAGAC AGGGCCTCTG GGTACCACTA CAACGCGCTC ACCTGTGAGG	480
GCTGCAAAGG TTTCTTCGGA AGAAGCATT CCAAGAACGC CGTGTACAAG TGTAAGAACG	540
GGGGCAACTG CGTGATGGAC ATGTACATGC GCAGGAAGTG CCAGGAGTGC CGGCTAAGGA	600
AGTGCAGAGA GATGGGGATG TTGGCTGAAT GTTTGTTAAC TGAAATCCAG TGTAATCTA	660
AACGGCTAAG GAAAAATGTG AAGCAGCAG CTGATCAGAC AGTGAATGAG GACGACAGCG	720
AAGGGCGTGA CTTGCGACAA GTGACCTCCA CAACCAAGTT TTGCAGGGAG AAAACGGAAC	780
TCACGGCAGA CCAGCAGACC CTCCTGGATT ATATTATGGA TTCGTACAAC AAACAGAGAA	840
TGCGTCAGGA AATCACAAAT AAAATCTTAA AAGAAGAATT TAGTGCAGAA GAAAATTTTC	900

TCATATTAAC AGAAATGGCA ACCAGCCATG TACAGATTCT CGTAGAATTC ACAAAAAAGC	360
TTCCAGGGTT TCAGACACTG GACCACGAAG ATCAGATTGC TTTGCTCAAA GGGTCCGCAG	1020
TGGAGGCCAT GTTTCTTCGT TCGGCGGAGA TTTTCAATAA GAAACTTCCT GCCGGACATG	1090
CAGACCTGTT GGAAGAAAGA ATTCGAAAGA GTGGTATCTC TGATGAGTAT ATAACCCCGA	1140
TGTTCACTTT CTATAAAAGT GTTGGAGAAC TCAAAATGAC TCAGGAGGAG TACGCTCTGC	1200
TCACAGCGAT CGTCATCTC TCTCCAGACA GACAATACAT CAAGGACAGA GAGGCGGTGG	1260
AGAAGCTGCA GGAGCCCCCTG CTTGATGTGC TACAAAAGCT GTGCAAGATG TACCAGCCTG	1320
AGAACCCACA GCATTTGCGC TGCCTCCTGG GTGCGCTGAC GGAAGTCCCG ACATTCAACC	1380
ATCACCACGC TGAGATGCTG ATGTCTTGGA GAGTGAATGA TCACAAGTTC ACCCCGCTCC	1440
TCTGTGAGAT CTGGGATGTG CAGTGATGGA CACCAGTGGG GCTGGCTCCT TGTCTCCTC	1500
GGAACAGAAA CCTTGTTTTG TTTGTACCTG GTTCACTCA AGAATCTCAA TGAATATTTA	1560
TGTGGCAATT ATACACCTCC CACGGTTGTA AATACAGACT AGATAGAAGT GCTTTCCCCA	1620
CACTGTATTT TACAAGGCTT CAGGAAACCC CACTGGCATG CCCTTTTGGC CTAATTAAAT	1680
CAATTGTTAC TTCAATTCTA TCTACTGAGC TAGGGGCATA TTATTCTTCA TTCGACAATA	1740
TTATATATAT TTTATAAAGT TGAGCTGTTT TCAACTGAGA CAATAAA	1787

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1960 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCCAGGGCAA CAGAGTCGGA GACCCCTGC CACCCCTC CCGATCGCCG GTGCAGTCAT	60
GAGCCCCGCC TCCCCCTGGT GCACGGAGAG GGGCGGGGCC TGGAAACAAGC AGGCTGCTTC	120
GTGACCCACT ATGTCTTCCC CCACAAGTTC TCTGGACACT CCCGTGCCTG GGAATGGTTC	180
TCCTCAGCCC AGTACCTCCG CCACGTCACC CACTATTAAG GAAGAGGGGC AGGAGACTGA	240
TCCTCCTCCA GGCTCTGAAG GGTCCAGCTC TGCTACATC GTGGTCATCT TAGAGCCAGA	300
GGATGAGCCT GAGCGCAAGC GGAAGAAGGG GCGGGCCCG AAGATGCTGG GCCATGAGCT	360
GTGCGCGGTG TGGGAGACA AGGCTTCGGG CTTCACCTAC AACGTGCTCA GCTGTGAAGG	420
GTGCAAAGGC TTCTTCGGC GCAGTGTGGT CCACGGTGGG GCGGGCGCT ATGCCTGTGG	480
GGGCAGCGGA ACCTGCCAGA TGGATGCCTT CATCGCGGC AAGTGCCAGC TGTGCGGCT	540

GCGCAAGTGC AAGGAGGCTG GCATGCGGGA GCAGTGGGTG CTCTCTGAGG AGCAGATTCTG	600
GAAGAAAAGG ATTCAGAAGC AGCAACAGCA GCAGCCACCA CCCCCATCTG AGCCAGCAGC	660
CAGCAGCTCA GGCCGGCCAG CGGCCCTCCCC TGGCACTTCG GAAGCAAGCA GCCAGGGCTC	720
CGGGGAAGGA GAGGGCATCC AGCTGACCGC GGCTCAGGAG CTGATGATCC AGCAGTTAGT	780
TGCCGCGCAG CTGCAGTGCA ACAAACGATC TTTCTCCGAC CAGCCCCAAG TCACGCCCTG	840
GCCCCTGGGT GCAGACCCTC AGTCCCGAGA TGCCCGTCAG CAACGCTTTG CCCACTTCAC	900
CGAGCTAGCC ATCATCTCGG TCCAGGAGAT TGTGGACTTT GCCAAGCAGG TGCCAGGGTT	960
CTTGCACTTG GGCCGGGAGG ACCAGATCGC CCTCCTGAAG GCGTCCACCA TTGAGATCAT	1020
GTTGCTACAA ACAGCCAGAC GGTACAACCA CGAGACAGAA TGCATCACGT TCCTGAAGGA	1080
CTTCACCTAC AGCAAGGACG ACTTCCACCG TGCAGGCTTG CAGGTGGAAT TCATCAATCC	1140
CATCTTCGAG TTCTCGCGGG CCATGCGGCG GCTGGGCCTG GACGATGCAG AGTATGCCTT	1200
GCTTATCGCC ATCAACATCT TCTCAGCCGA TCGGCCTAAT GTGCAGGAGC CCAGCCGTGT	1260
GGAGGCCCTG CAGCAGCCCT ACGTGGAGGC GCTCCTCTCC TACACGAGGA TCAAGCGCCC	1320
ACAGGACCAG CTCCGCTTCC CACGCATGCT CATGAAGCTG GTGAGCCTGC GCACCCTCAG	1380
CTCCGTGCAC TCGGAGCAGG TCTTTGCATT GCGACTCCAG GACAAGAAGC TGCCGCCCTT	1440
GCTGTCCGAG ATCTGGGATG TGCACGAGTA GGGGCAGCCA CAAGTGCCCC AGCCTTGGTG	1500
GTGTCTTCTT GAAGATGGAC TCTTCACCTC TCCTCCTGGG GTGGGAGGAC ATTGTCACGG	1560
CCCAGTCCCT CGGGCTCAGC CTCAAACTCA GCGGCAGTTG GCACTAAGAA GGCCCCACCC	1620
CACCCATTGA GTCTTCCAAG AGTGGTGAGG GTCACAGGTC CTAGCCTCTG ACCGTTCCCA	1680
GCTGCCCTCC CACCCACGCT TACACCTCAG CCTACCACAC CATGCACCTT GAGTGGAGAG	1740
AGGTTAGGGC AGGTGGCCCC CCACAGTTGG GAGACCACAG GCCCTCTCTT CTGCCCCTTT	1800
TATTTAATAA AAAAACAAAA ATAAAGTTTG AGTACAAGCC AAAAAAAAAA AAAAAAAAAA	1860

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2389 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TCCTCAAGTT GTGGCTTTCA GAGGAAGGAT CTGGAAACAA GCAGTTGTGT CAGTATAAAG	60
AAGAAGCGTA GACTTGAGGA CTTACTCATA GTGATATCCG ATAGCGATGG AGAGGAAACA	120

AAAGAGGAGA ATGGATTGCA GAAAACGAAG AAAAAACAGT CGAACAGATC AAAGTGTTTG	130
GCTAAAAGAA AAGTTGCACA CATGTCAGAA GAAGAACAAT TTGCTTTGGC TCTCAAGATG	240
AGTGAGCAGG AAGCTAGGGA GGTGAATAAC CAGGAGGAGA AAGAAGAGGA GCTCTTGCGG	300
AAAGCCATTG CTGAAAGCCT GAATAGTTGC TGGTCTTCTG CTGCTTCTGC TACCAGATCT	360
CGACCTCTGG CTGCTGAACT ATCTTCACAT TCCCATCAAG AGAACACCAA AGACTCTGGG	420
ACCACTGAAG GCGTATGGCA GCTGGTACCT CCATCACTGT GTAAAGGCTC ACATGTCAGT	480
CAGGGAAACG AGGCTGAGCA AAGAAAGGAG CCCTGGGACC ACAATGAAAA CACTGAAGAG	540
GAGCCGGTCT CTGGCAGCTC AGGAAGCTGG GACCAGTCAA GCCAGCCAGT GTTTGAGAAT	600
GAGAACGTTA AATGTTTTGA CAGATGTACT GGCCACTTGG CTGAGCACAC ACAGTGTTGG	660
AAGCCACAGG AAAGTACTGG GAGTGGTTAT GCTTTTTCCA AAGCTGTCCA GGGTAGGGGG	720
GACACGTCTA GGCAATGCCT TCCTATCCCA GCAGACACAA AAGGTCTCCA GGACACTGGG	780
GGCACTGTGC ACTACTACTG GGGTATTCCA TTCTGCCCTG CTGGAGTAGA TCCCAATCAA	840
TACACCAATG TCATTCTCTG CCAGTTAGAG GTTTATCAGA AGAGCCTGAA AATGGCTCAG	900
AGACAGCTTG TTAAAAAAG AGGGTTTGGG GAACCAGTGT TACCTAGACC TCCTTTTCTG	960
ATCCAGAATG AATGTGGCCA AGAAGATCAG ACTAGTGACA AAAATGAAGG CATCTCAGAA	1020
GATATGGGAG ATGAAGCCAA AGAGGAAAGG CAGGAATCTA GGGCATCTGT CTGGCACTCA	1080
GAAACCAAGG ATTTTCAAAA AAGTCCAATT AAAAGCTTGA AACAGAAACT TTTGTTGGAG	1140
GAAGAACCAA CAACCAGTCG TGGTCAGTCT TCCCAAGGTC TGTTTGTTGA AGAAACCTCT	1200
GAAGAAGGTC TGAAGAGTTC GGAAGGAGAC AACTCTGTGC CCACCACGCA AAGCATTGCA	1260
GCTTTGACCA GTAAGAGAAG TTTAGTTCTT ATGCCGAAA GTTCTGCAGA AGAAATCACT	1320
GTTTGCCCTG AGACACAGTT AAGTTTCCTT GAACCCCTTG ACCTCAATAG AGAAGACTCT	1380
CCAGATAGCA GAGAGCTCCC CATTGAAGTA AGGATGGCAG TGGGCGATAA GCAGGTTGCT	1440
AATAGGGAAG ATTGTATGAA GGAAAACCTT CCTCCTGCAG TCTCATCTAG TACCCGGGTA	1500
TCCTGCCCCAC TGTGTAACCA AGACTTTCCT CCCACAAAGA TTGAACAGCA TGCCATGTAC	1560
TGCAATGGTC TGATGGAGCA GGAAACAGTG TTGACTCGGA GACGAAGAGA GGCCAAGAAC	1620
AAGAGTGACG GTCGGACAGC TGCACAGCCG GCTCTGGATG CCAACAGGAA GGAGAAGTGT	1680
TATCTATGTA AGTCCCTGGT TCCACTTGGG GAGTATCAGT GCCATGTGGA GGCTGTCTC	1740
CAGCTTGCAA AGGTTGACAG AGAAGATGGG ATTGAAGGGA CAAGGAGACC AAGGGTGTGT	1800
GCACCTGTGG AGGGGAAACA ACAGCAGCGG CTGAAGAAAGT CAAAGGACAA AGGCCATAGT	1860
CAAGGCCGAC TCCTCAGTCT CTTGGAGCAG TCTGAGCATA GGACCACAGG TGTAGAGAAA	1920
AAACCCAAGT ATTCGGAAGT AAGAACCCTC AGGATGCCCT CACCAGAGGT GGAAGAGGCT	1980

AGCTGCAGCA GAGAGATGCA GAGTACCCTC TCACAGCTCA ACTTAAATGA GTCTCCCATC	2040
AAGTCTTTTG TTCCTGTTTC AGAAGCTACA AATTGCTTAG TGGACTTTAA AGAACAGTTT	2100
GCTTTCCGGT CACGAACTAA ATCAGGCAGG GAAAGGAGGA GAAAATCTTG AATTTCTTGA	2160
GACTGGAAGG TTGACCAGAA CACACATCGT TGGGTTGATC GTGTTCATTA AGTATAGTGG	2220
TCTCTAGTTT GTGGTGAGAG TTCTGACCCT GTTGTTATCA CCACCAGCAC CCATTTCAGTA	2280
TCCTGGCTTT ATATTTTATA AGATCAGTTC AGACAACTGT GAATATTATT CTGTTTGAAT	2340
TTGCTTATAG TTAAAATTTA AATATATTTA TCTTTGTATG AAAAAAAAAA	2389

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1922 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GGATGCCCTG GCTGCTCTTG TGGATGCTGC AGCTTCTGCA CCCCAGATGG ATGTTTCCAA	60
AACAAAAGAG AGTAAGCATG AAGCTGCCAG GTTAGAAGAA AATTGAGAA GCAGGTCAGC	120
AGCAGTTAGT GAACAGCAGC AGCTAGAGCA GAAAAACCTG GAGGTGGAGA AGAGATCTGT	180
TCAGTGTGTG TGCATTCTT CAGCCCTTCC AAGTGGCAAG GCCCAGCCTC ATGCCTCAGT	240
AGTGTATTCT GAGGCTGGGA AAGATAAAGG GCCTCCTCCA AAATCCAGAT ATGAGGAAGA	300
GCTAAGGACC CGAGGGAAGA CTACCATTAC TGCAGCTAAC TTCATAGACG TGACCATCAC	360
CCGGCAAATT GCCTCGGACA AGGATGCGAG GGAACGTGGC TCTCAAAGTT CAGACTCTTC	420
TAGTAGCTTG TCTTCTCACA GGTATGAAAC GGCTAGTGAT GCCATTGAGG TGATAAGTCC	480
CGCCAGCTCA CCTGCACCAC CCCAGGAAAA GCCACAGGCC TATCAGCCAG ACATGGTTAA	540
GGCAAATCAA GCAGAAAATG AGTCCACTCG ACAGTATGAA GGTCCACTGC ATCATTATCG	600
GTCCCAGCAG GAATCACCAT CTCCACAGCA ACAGCCACCA CTGCCCCCAT CTTCCCAGTC	660
AGAGGGAATG GGACAGGTGC CCAGGACCCA TCGACTGATC ACACTTGCTG ACCACATCTG	720
TCAAATTATC ACACAAGATT TTGCTAGAAA TCAAGTTCCC TCGCAGCCTT CTACTTCTAC	780
ATTCCAAACT TCACCATCTG CTTTGTGATC CACACCTGTA AGAACTAAAA CCTCAAGCCG	840
CTACAGCCCA GAATCACAGT CTCAGACTGT CTTGCATCCC AGACCAGGTC CTAGAGTCTC	900
TCCAGAAAAT CTTGTGGATA AATCCCGGGG AAGCAGGCCT GGAAAATCTC CAGAGAGGAG	960
TCATATCCCA TCAGAGCCCT ATGAGCCCAT CTCCCACCC CAAGGCCCTG CTGTGCATGA	1020

GAAGCAGGAC AGCATGTTGC TCTTGTCA	GAGGGGAGTG GACCCTGCTG AGCAAAGGAG	1080
TGATTCTCGA TCACCAGGAA GTATAAGCTA	CTTGCCCTTCA TTCTTCACCA AGCTTGAAAG	1140
CACATCACCC ATGGTTAAAT CAAAGAAACA	GGAAATTTTT CGTAAGTTGA ACTCTTCTGG	1200
TGGAGGTGAC TCTGATATGG CAGCTGCTCA	GCCAGGAACA GAGATCTTCA ATCTGCCAGC	1250
AGTTACCACA TCAGGTGCAG TGAGCTCAAG	AAGCCATTCT TTTGCTGATC CCGCCAGTAA	1320
CCTTGGTCTA GAAGACATCA TCAGAAAGGC	TCTCATGGGA AGTTTTGATG ATAAAGTTGA	1380
AGATCATGGT GTTGTCTATGT CCCATCCTGT	GGGCATTATG CCTGGTAGTG CCAGCACCTC	1440
AGTGGTGACG AGCAGCGAGG CACGGAGAGA	TGAAGGGGAG CCATCACCTC ATGCAGGAGT	1500
ATGCAAACCA AAGCTGATCA ACAAATCAAA	CAGCAGGAAG TCTAAATCTC CTATTCCTGG	1560
GCAAAGCTAT TTAGGAAGCTG AAAGGCCTTC	TTCTGTCTCC TCTGTGCATT CAGAAGGTGA	1620
TTACCACAGG CAGACACCAG GATGGGCATG	GGAAGATCGG CCCTCTTCAA CAGGTTCTAC	1680
TCAGTTCCCT TACAACCCTC TGACCATACG	GATGCTCAGC AGTACACCAC CTACACAGAT	1740
CGCATGCGCC CCATCTGCCA TCACCCAAGC	AGCTCCACAT CAACAGAACC GCATCTGGGA	1800
GAGGGAGCCT GCCCCGCTCC TCTCAGCGCA	GTATGAGACA CTGTCTGATA GTGACGACTG	1860
AGCTGTGCGT GGGAGAGCGC TCTGGCTTTG	GTTTTTATTG AAGATTTAAA AAAAAAAAAA	1920
AA		1922

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 6 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GURAGU

6

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:
GATCCGGGTA GGGTTCACCG AAAGTTCAC T CGA 33

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:
CTAGACAAGG GTTCAATGCA CTTGTCCATC G 31

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
AGGTCA 6

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1677 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGCTAACAG AACACGCGGC AGGCCCTCTG GGGCAGAATC TGGATTGGA ATCGTACTCC	60
CCATACAACA ATGTCCCGTT TCCTCAAGTT CAGCCACAGA TTTCCTCCTC GTCTTACTAT	120
TCCAACCTGG GCTTCTACCC CCAACAACCG GAAGACTGGT ATTCTCCTGG CATCTATGAA	180
CTCAGGCGAA TGCCCGCTGA GACTGGGTAC CAGGGAGAGA CTGAGGTATC AGAGATGCCT	240
GTGACAAAGA AGCCGCGAAT GGCCGCGGCA TCGGCAGGCA GAATAAAAGG GGATGAGCTG	300
TGTGTTGTCT GTGGAGACAG GGCCTCTGGG TACCACTACA ACGCGCTCAC CTGTGAGGGC	360

TGCAAAGGTT TCTTCCGAAG AAGCATTACC AAGAACGCCG TGTACAAGTG TAAGAACGGG	420
GGCAACTGCG TGATGGACAT GTACATGCGC AGGAAGTGCC AGGAGTGCCG GCTAAGGAAG	430
TGCAGAGAGA TGGGGATGTT GGCTGAATGT ATGTATACAG GTTTGTTAAC TGAAATCCAG	540
TGTAAATCTA AACGGCTAAG GAAAAATGTG AAGCAGCACG CTGATCAGAC AGTGAATGAG	600
GACGACAGCG AAGGGCGTGA CTTGCGACAA GTGACCTCCA CAACCAAGTT TTGCAGGGAG	660
AAAACGGAAC TCACGGCAGA CCAGCAGACC CTCCTGGATT ATATTATGGA TTCGTACAAC	720
AAACAGAGAA TGCCTCAGGA AATCACAAAT AAAATCTTAA AAGAAGAATT TAGTGCAGAA	780
GAAAATTTTC TCATATTAAC AGAAATGGCA ACCAGCCATG TACAGATTCT CGTAGAATTC	840
ACAAAAAAGC TTCCAGGGTT TCAGACACTG GACCACGAAG ATCAGATTGC TTTGCTCAAA	900
GGGTCCGCAG TGGAGGCCAT GTTTCTTCGT TCGGCGGAGA TTTTCAATAA GAAACTTCCT	960
CCCGGACATG CAGACCTGTT GGAAGAAAGA ATTCGAAAGA GTGGTATCTC TGATGAGTAT	1020
ATAACCCCGA TGTTCAGTTT CTATAAAAGT GTTGGAGAAC TCAAATGAC TCAGGAGGAG	1080
TACGCTCTGC TCACAGCGAT CGTCATCCTC TCTCCAGACA GACAATACAT CAAGGACAGA	1140
GAGGCGGTGG AGAAGCTGCA GGAGCCCCTG CTTGATGTGC TACAAAAGCT GTGCAAGATG	1200
TACCAGCCTG AGAACCACACA GCATTTCGCC TGCCTCCTGG GTCGCCTGAC GGAACTCCGG	1260
ACATTCAACC ATCACCACGC TGAGATGCTG ATGTCTTGGA GAGTGAATGA TCACAAGTTC	1320
ACCCCGCTCC TCTGTGAGAT CTGGGATGTG CAGTGATGGA CACCAGTGGG GCTGGCTCCT	1380
TGTCTCCTC GGAACAGAAA CCTTGTTTCG TTTGTACCTG GTTTCACTCA AGAATCTCAA	1440
TGAATATTTA TGTGGCAATT ATACACCTCC CACGGTTGTA AATACAGACT AGATAGAACT	1500
GCTTTCCCCA CACTGTATTT TACAAGGCTT CAGGAAACCC CACTGGCATG CCCTTTTGGC	1560
CTAATTAAAT CAATTGTTAC TTCAATTCTA TCTACTGAGC TAGGGGCATA TTATTCTTCA	1620
TTGACAATA TTATATATAT TTTATAAAGT TGAGCTGTTT TCAACTGAGA CAATAAA	1677

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 60 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TGGTCACCCA GGCTTCTGCT TCAGTCTCTC CTCCTTCTCC TCCTCAGCCC ACTGTCTCCT	60
---	----



(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

AAAATTACTG GGCACTAGAA AGGAAGACTG GGCTCCGAAT CCTCTTAGAG CCTTGGACAT 60

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 246 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGAGAAGAAC CGAGTTCTGA GAGTCTACAG CAAAATTACT GGGCACTAGA AAGGAAGACT 60
GGGCTCCGAA TCCTCTTAGA GCCTTGGACA TCTCTGGCCC AAAGCAATCC AAGGATCTTA 120
TTTGAGGACC ACCATCCCAG AAGTACTTTC TCAAGGTTGA AAAGTTGGAG TGGTAGCCAA 180
GATGAATCTG ATTGGGCACT CCATTTACAG GCTACGGACG AGTTTTCTCT TTCTGAAAGC 240
TTATTT 246